Write your name here

Surname	Other names

Pearson Edexcel International Advanced Level

Centre Number | Candidate Number

Statistics S1

Advanced/Advanced Subsidiary

Wednesday 13 June 2018 – Morning
Time: 1 hour 30 minutes

Paper Reference
WST01/01

You must have:
Mathematical Formulae and Statistical Tables (Blue)

Total Marks

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B). Coloured pencils and highlighter pens must not be used.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information

- The total mark for this paper is 75.
- The marks for **each** question are shown in brackets – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over ▶

©2018 Pearson Education Ltd.

P 5 0 8 0 7 A 0 1 2 8

Pearson

https://xtremepape.rs/

Leave blank

1. A random sample of 10 cars of different makes and sizes is taken and the published miles per gallon, p, and the actual miles per gallon, m, are recorded. The data are coded using variables $x = \frac{p}{10}$ and $y = m - 25$

The results for the coded data are summarised below.

x	6.89	3.67	5.92	5.04	4.87	3.92	4.71	5.14	3.65	5.23
y	30	3	22	15	13	8	15	13.5	3	19

(You may use $\sum y^2 = 2628.25$ $\sum xy = 768.58$ $S_{xx} = 9.259\,24$ $S_{xy} = 74.664$)

(a) Show that $S_{yy} = 626.025$

(2)

(b) Find the product moment correlation coefficient between x and y.

(2)

(c) Give a reason to support fitting a regression model of the form $y = a + bx$ to these data.

(1)

(d) Find the equation of the regression line of y on x, giving your answer in the form $y = a + bx$.
Give the value of a and the value of b to 3 significant figures.

(3)

A car's published miles per gallon is 44

(e) Estimate the actual miles per gallon for this particular car.

(3)

(f) Comment on the reliability of your estimate in part (e). Give a reason for your answer.

(2)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

P 5 0 8 0 7 A 0 2 2 8
https://xtremepape.rs/

Leave blank

Question 1 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 1 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 1 continued

(Total 13 marks)

Q1

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

2. Two youth clubs, *Eastyou* and *Westyou*, decided to raise money for charity by running a 5 km race. All the members of the youth clubs took part and the time, in minutes, taken for each member to run the 5 km was recorded.

The times for the *Westyou* members are summarised in Figure 1.

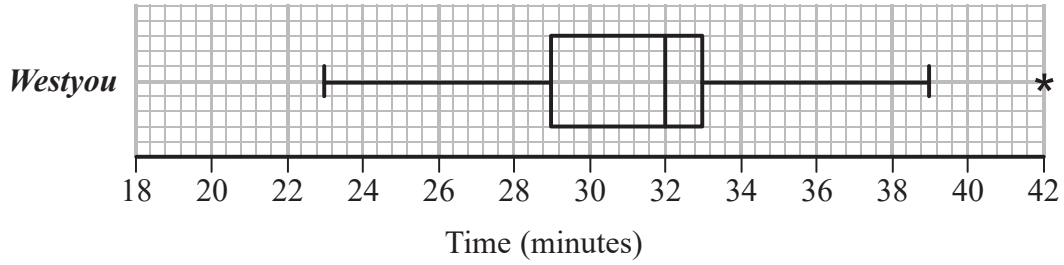


Figure 1

(a) Write down the time that is exceeded by 75% of *Westyou* members.

(1)

The times for the *Eastyou* members are summarised by the stem and leaf diagram below.

Key: 2|0 means 20 minutes

Stem	Leaf	
2	0 2 3 4	(4)
2	5 6 8 8 8 9 9	(7)
3	0 0 0 0 0 1 1 1 2 2 2 2 3 4	(14)
3	5 5 5 7 9	(5)

(b) Find the value of the median and interquartile range for the *Eastyou* members.

(3)

An outlier is a value that falls either

more than $1.5 \times (Q_3 - Q_1)$ above Q_3

or more than $1.5 \times (Q_3 - Q_1)$ below Q_1

(c) On the grid on page 7, draw a box plot to represent the times of the *Eastyou* members.

(4)

(d) State the skewness of each distribution. Give reasons for your answers.

(3)

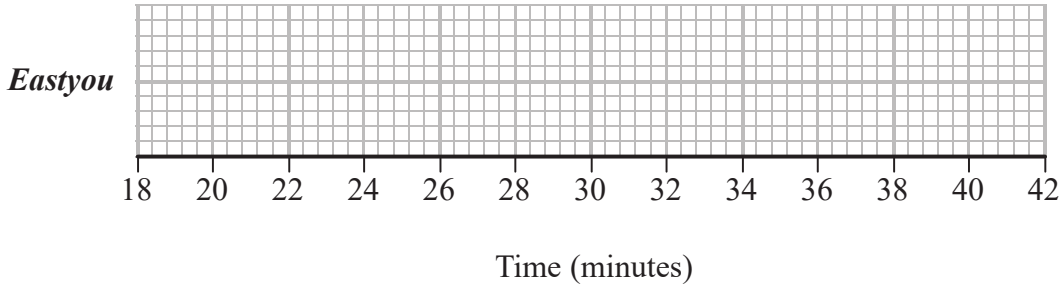
P 5 0 8 0 7 A 0 6 2 8

https://xtremepape.rs/

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 2 continued



Turn over for a spare grid if you need to redraw your box plot.

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 2 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

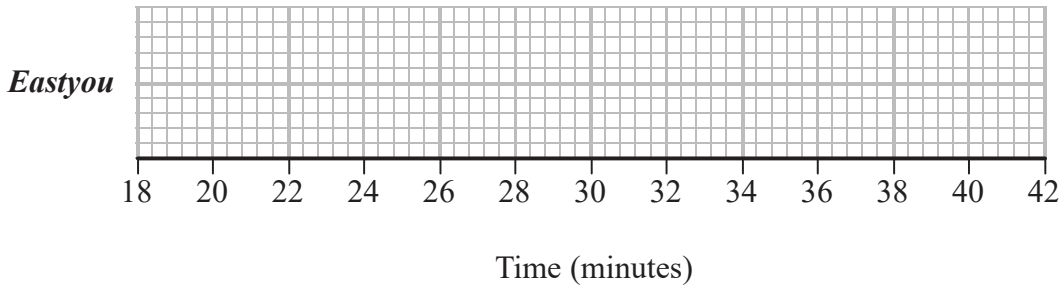
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 2 continued

Only use this grid if you need to redraw your box plot.



Q2

(Total 11 marks)

P 5 0 8 0 7 A 0 9 2 8

https://xtremepape.rs/

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

3. A manufacturer of electric generators buys engines for its generators from three companies, R, S and T.

Company R supplies 40% of the engines.

Company S supplies 25% of the engines.

The rest of the engines are supplied by company T.

It is known that 2% of the engines supplied by company R are faulty, 1% of the engines supplied by company S are faulty and 2% of the engines supplied by company T are faulty.

An engine is chosen at random.

(a) Draw a tree diagram to show all the possible outcomes and the associated probabilities. **(2)**

(b) Calculate the probability that the engine is from company R and is not faulty. **(1)**

(c) Calculate the probability that the engine is faulty. **(2)**

Given that the engine is faulty,

(d) find the probability that the engine did **not** come from company S. **(2)**

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 3 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 3 continued

(Total 7 marks)

Q3

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

4. A discrete random variable X has probability function

$$\mathrm{P}(X = x) = \begin{cases} k(2 - x) & x = 0, 1 \\ k(3 - x) & x = 2, 3 \\ k(x + 1) & x = 4 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

(a) Show that $k = \frac{1}{9}$ **(2)**

Find the exact value of

(b) $\mathrm{P}(1 \leqslant X < 4)$ **(1)**

(c) $\mathrm{E}(X)$ **(2)**

(d) $\mathrm{E}(X^2)$ **(2)**

(e) $\mathrm{Var}(3X + 1)$ **(3)**

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

P 5 0 8 0 7 A 0 1 4 2 8
https://xtremepape.rs/

Leave blank

Question 4 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 4 continued

(Total 10 marks)

Q4

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

5. The weights, in grams, of a random sample of 48 broad beans are summarised in the table.

Weight in grams (x)	Frequency (f)	Class midpoint (y)
$0.9 < x \leqslant 1.1$	9	1.0
$1.1 < x \leqslant 1.3$	12	1.2
$1.3 < x \leqslant 1.5$	11	1.4
$1.5 < x \leqslant 1.7$	8	1.6
$1.7 < x \leqslant 1.9$	3	1.8
$1.9 < x \leqslant 2.1$	3	2.0
$2.1 < x \leqslant 2.7$	2	2.4

(You may assume $\sum fy^2 = 101.56$)

A histogram was drawn to represent these data. The $2.1 < x \leqslant 2.7$ class was represented by a bar of width 1.5 cm and height 1 cm.

(a) Find the width and height of the $0.9 < x \leqslant 1.1$ class.
(3)

(b) Give a reason to justify the use of a histogram to represent these data.
(1)

(c) Estimate the mean and the standard deviation of the weights of these broad beans.
(4)

(d) Use linear interpolation to estimate the median of the weights of these broad beans.
(2)

One of these broad beans is selected at random.

(e) Estimate the probability that its weight lies between 1.1 grams and 1.6 grams.
(1)

One of these broad beans having a recorded weight of 0.95 grams was incorrectly weighed. The correct weight is 1.4 grams.

(f) State, giving a reason, the effect this would have on your answers to part (c). Do not carry out any further calculations.
(2)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

P 5 0 8 0 7 A 0 1 8 2 8
https://xtremepape.rs/

Leave blank

Question 5 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 5 continued

(Total 13 marks)

Q5

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

6. The waiting time, L minutes, to see a doctor at a health centre is normally distributed with $L \sim \mathrm{N}(\mu, \sigma^2)$.

Given that $\mathrm{P}(L < 15) = 0.9$ and $\mathrm{P}(L < 5) = 0.05$

(a) find the value of μ and the value of σ.

(6)

There are 23 people waiting to see a doctor at the health centre.

(b) Determine the expected number of these people who will have a waiting time of more than 12 minutes.

(3)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

P 5 0 8 0 7 A 0 2 2 2 8
https://xtremepape.rs/

Leave blank

Question 6 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 6 continued

(Total 9 marks)

Q6

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

7. Events A and B are such that

$$\mathrm{P}(A) = 0.5 \qquad \mathrm{P}(A \mid B) = \frac{2}{3} \qquad \mathrm{P}(A' \cup B') = 0.6$$

(a) Find $\mathrm{P}(B)$

(3)

(b) Find $\mathrm{P}(A' \mid B')$

(2)

The event C has $\mathrm{P}(C) = 0.15$

The events A and C are mutually exclusive.

The events B and C are independent.

(c) Find $\mathrm{P}(B \cap C)$

(2)

(d) Draw a Venn diagram to illustrate the events A, B and C and the probabilities for each region.

(5)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

P 5 0 8 0 7 A 0 2 6 2 8
https://xtremepape.rs/

Leave blank

Question 7 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

https://xtremepape.rs/

Leave blank

Question 7 continued

Q7

(Total 12 marks)

TOTAL FOR PAPER: 75 MARKS

END

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

https://xtremepape.rs/